

Genome sequences from the sea

Jed Fuhrman

Despite their diminutive stature, phytoplankton have a huge global influence. The genomes of four strains of phytoplankton have now been completely sequenced, revealing their genetic adaptations to distinct marine niches.

With the sequencing of microbial genomes now almost routine in some circles, one could be forgiven for feeling a little jaded. It is possible to lose sight of just how much can be learned from such an exercise. But the reports by Rocap *et al.*¹ and Palenik *et al.*² in this issue, along with a paper by Dufresne *et al.*³ in *Proceedings of the National Academy of Sciences*, powerfully demonstrate how genomic studies can lead to a new understanding of biodiversity, ecology, biological efficiency and biogeochemistry.

About half of global photosynthesis and oxygen production is accomplished by single-celled planktonic organisms (phytoplankton) that live in the top layer of the ocean, where enough light penetrates to support their growth. The most plentiful of these are the cyanobacteria — tiny, chlorophyll-containing phytoplankton that have no membrane-bound nucleus. There are two basic types. The *Synechococcus* strains have a diameter of about 0.9 μm and were discovered to be abundant in seawater in 1979 (ref. 4). With a diameter of about 0.6 μm , meanwhile, the *Prochlorococcus* strains (Fig. 1) are the smallest of all the phytoplankton; they are also the most abundant, yet were discovered only 15 years ago⁵.

Now the complete genomes of three strains of *Prochlorococcus*^{1,3} and one strain of *Synechococcus*² have been sequenced and analysed. The results are remarkable for what they show not only about the differences between these close relatives, but also about

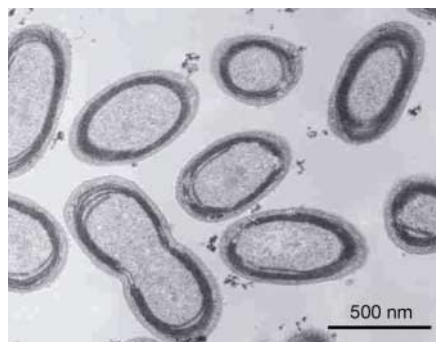


Figure 1 ***Prochlorococcus***: the smallest of the marine phytoplankton, and the genus with the smallest known genomes of any oxygen-generating photosynthetic organism. This electron micrograph shows strain MIT9313, one of the strains whose complete genome sequence is now reported¹⁻³.

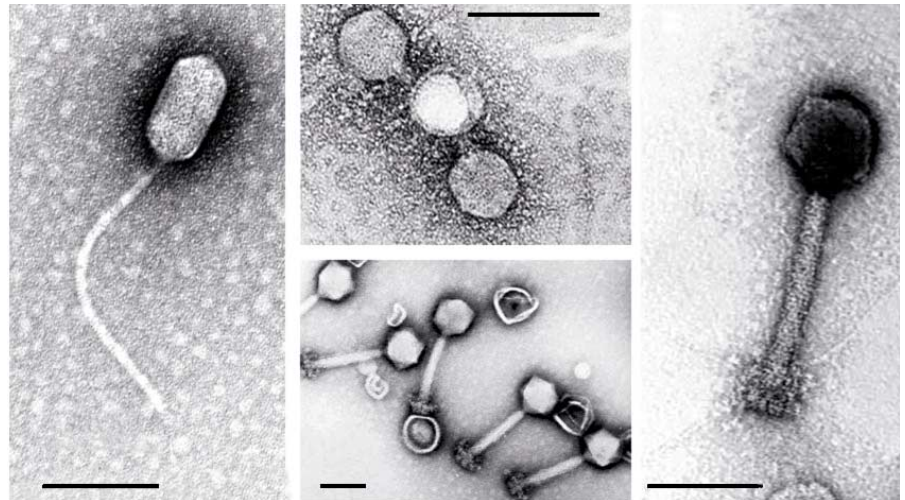


Figure 2 **Marine cyanophages**. Cyanophages are viruses that infect cyanobacteria; those shown here infect *Prochlorococcus*. The newly reported whole-genome sequence data¹⁻³ suggest that cyanophages might be an important factor driving the evolution of cyanobacteria. Also in this issue, Sullivan *et al.*¹¹ describe the first isolation of *Prochlorococcus*-infecting cyanophages, with some phages being capable of infecting multiple strains of *Prochlorococcus*, and even of crossing to the closely related *Synechococcus*. Cross-infection is a possible mechanism of gene transfer between different hosts. Scale bars represent 100 nm. Left, siphovirus P-SS2; top, podovirus P-SSP9; right and bottom, myoviruses P-RSM2.

the extent to which some strains economize on DNA. Indeed, as Rocap *et al.*¹ and Dufresne *et al.*³ describe, two of the *Prochlorococcus* strains have genomes so small — about 1.7 million base pairs — that they might represent the minimal genome of an oxygen-generating, photosynthetic organism.

Why economize on DNA? A genome represents the complete genetic repertoire of an organism: it contains specifications for all the machinery needed to create the organism and to regulate its operation, growth and reproduction. Control of genome size involves a trade-off between efficiency and versatility. A small genome reduces the amount of extra 'baggage' that must be maintained and propagated, but also limits an organism's ability to exploit different resources. Larger genomes provide for this possibility, and also permit back-up should a gene be damaged or lost. But they require more material and energy to maintain.

For some *Prochlorococcus* strains, it seems that the minimalist approach to genomes works best^{1,3}. *Prochlorococcus* is a photosynthetic 'autotroph' (an organism that uses inorganic sources of carbon), and its genome must encode proteins that allow all cellular components to be synthesized from simple

inorganic compounds, with power for all this coming from sunlight. In the top layer of the sea, nutrients such as nitrogen and phosphorus are often extremely dilute. So the two-pronged strategy of being physically small (reducing cellular requirements) and of economizing on DNA (which contains both nitrogen and phosphorus) can clearly be very successful — although there are other strategies, of course.

The tiny genomes now studied belong to the high-light-adapted *Prochlorococcus* strain MED4 (ref. 1) and the very-low-light-adapted strain SS120 (ref. 3), which might be able to get away with such genomic paucity in part because of the stability of their environments. They grow optimally in the bright (near-surface) and dim (deeper) portions of the water column, respectively; SS120 cannot photosynthesize at the light levels that are optimal for MED4 (ref. 6). The different kinds of *Prochlorococcus* can do well when specialized for certain depths because the warm ocean water where they are typically found is usually stratified by density — so different depths rarely mix with each other. Resources such as light and nutrients are also stratified (see Fig. 1 of ref. 1, page 1043).

For the other strains now studied, larger genome size probably relates to the use of

a broader variety of resources. The third *Prochlorococcus* strain, MIT9313, has adaptations that seem to place its optimum depth between that of MED4 and SS120 in the water column. Rocap *et al.* find that it also has a larger genome (2.4 million base pairs) and more genomic versatility, perhaps enabling it to exploit the diverse resources available in the 'transition' zone where it lives. According to Palenik *et al.*², the *Synechococcus* strain WH8102 also has a larger genome (again, about 2.4 million base pairs). This organism is known to fare particularly well under conditions of upwelling and vertical mixing, where individual cells can be exposed to a large variety of nutrient and light conditions. Its bigger genome and greater number of genes presumably give it the ability to acclimate. For instance, the sequences imply that *Synechococcus* strain WH8102 can use ammonium, nitrite, nitrate, urea, cyanate, amino acids and peptides as sources of nitrogen. *Prochlorococcus* MIT9313 can use all except nitrate and cyanate, MED4 only ammonium, urea, cyanate and peptides, and SS120 only ammonium and amino acids.

The remarkable partitioning of resources such as light and nutrients between these close cyanobacterial relatives beautifully illustrates the question of the meaning and significance of the term 'biodiversity' when it comes to bacteria. Although animal and plant diversity is continually in the news, with evidence mounting about serious losses, microbiologists are often hung up on simple questions such as how to measure diversity or define a species. Because most bacteria lack taxonomically useful morphological features, classification studies now rely heavily on molecular sequences. New molecular biological techniques allow microbial diversity to be determined directly from natural communities, without needing to cultivate bacteria in the laboratory (which is often problematic)⁷.

But it is not easy to interpret the results of sequence-based diversity studies in a classical ecological framework, or to reconcile the results obtained by different methods. For example, one recent study used a bioinformatic analysis of sequence databases to estimate that marine bacterial species number only a few thousand⁸. In contrast, another used abundance curves to calculate that there may be as many as two million distinct taxa⁹. Interestingly, both sorts of analysis would have lumped most of the cyanobacteria sequenced here¹⁻³ into a single taxon or perhaps a single species, because such analyses tend to use a criterion of more than 97% sequence identity in the small subunit of ribosomal RNA to define a species or taxon (and the three *Prochlorococcus* strains fit that criterion). Clearly that would be a mistake, given what we know about the physiological and, now, the genomic differences. And that

suggests that previous estimates of the total diversity of marine bacteria are rather low.

Can genomes provide clues to how biodiversity is created or maintained? Indeed they can: the new studies¹⁻³ tell a story of rapid evolutionary adaptation, including gene loss and a significant transfer of genes from other bacteria. Of particular note are changes in genes controlling the cell surface; these affect susceptibility to viruses (Fig. 2), and possibly recognition by protists that feed on the cyanobacteria — which is consistent with ecological investigations that implicate viruses or grazers as important selective agents in bacterial evolution¹⁰. Ironically, several of the gene-transfer events might themselves have been mediated by viruses, as demonstrated by the flanking of transferred genetic regions by viral-like integrase sequences²; integrases are enzymes that insert viral nucleic acids into a host genome.

Yet more surprises hide in these data. For example, as mentioned above, *Prochlorococcus* MED4 maintains genes that allow cyanate to be used as a nitrogen source. But until now, biological oceanographers have not thought of cyanate as being a nitrogen source of any significance. The same applies to phosphonates (molecules with C-P bonds) as a source of phosphorus. Then there is the mystery of how these organisms can perform some enzymatic functions even

though they apparently lack genes similar to those that perform these functions in other organisms. For instance, there is no recognizable gene for carbonic anhydrase, which converts bicarbonate to CO₂ for photosynthesis. In fact, about one-third of each of the newly sequenced genomes bears no relation to anything previously identified. These genes are a veritable treasure-trove that should point to novel physiological and ecological phenomena. Who would have thought 20 years ago that marine ecologists would be eagerly awaiting gene sequences? ■

Jed Fuhrman is in the Department of Biological Sciences and the Wrigley Institute for Environmental Studies, University of Southern California, Los Angeles, California 90089-0371, USA.
e-mail: fuhrman@usc.edu

1. Rocap, G. *et al.* *Nature* **424**, 1042–1047 (2003).
2. Palenik, B. *et al.* *Nature* **424**, 1037–1042 (2003).
3. Dufresne, A. *et al.* *Proc. Natl Acad. Sci. USA*; published online 13 August 2003 (doi: 10.1073/pnas.1733211100).
4. Waterbury, J. B., Watson, S. W., Guillard, R. L. L. & Brand, L. E. *Nature* **227**, 293–294 (1979).
5. Chisholm, S. W. *et al.* *Nature* **334**, 340–343 (1988).
6. Moore, L. R. & Chisholm, S. W. *Limnol. Oceanogr.* **44**, 628–638 (1999).
7. Pace, N. R. *Science* **276**, 734–740 (1997).
8. Hagstrom, A. *et al.* *Appl. Env. Microbiol.* **68**, 3628–3633 (2002).
9. Curtis, T. P., Sloan, W. T. & Scannell, J. W. *Proc. Natl Acad. Sci. USA* **99**, 10494–10499 (2002).
10. Fuhrman, J. A. *Nature* **399**, 541–548 (1999).
11. Sullivan, M. B., Waterbury, J. B. & Chisholm, S. W. *Nature* **424**, 1047–1051 (2003).

Turbulence

Suddenly it's chaos

George Pickett

Injecting vortices into a rotating sample of superfluid helium-3 shows a sudden switch from smooth to chaotic behaviour, and throws light on turbulence — one of the last unsolved problems of classical physics.

Leonardo sketched it. Hydrodynamicists model it. Turbulence in fluid flows is universal. Not only does it impinge directly on human existence, but it affects physical behaviour on all scales, from the dynamics of galaxies to particle motion inside the atomic nucleus. Turbulence is clearly important and we know a great deal about it, but there is no satisfactory theory for it. One way forward is to attack the problem in its simplest form, as A. P. Finne *et al.* have done in their experiments reported in this issue (*Nature* **424**, 1022–1025; 2003).

This group is studying vortices in the simplest of fluids, superfluid helium. In a normal liquid the constituent particles are distributed over many possible quantum-mechanical energy states. Constant collisions ensure that the behaviour on a microscopic level is a frantic redistribution of particles among the available states. To model that, we would need one set of

equations to track the career of every constituent particle, and that is quite beyond us. In contrast, when liquid helium is cooled to temperatures close to absolute zero, a large fraction of the constituent particles occupy the lowest available energy state and so they all behave alike — they march in step, rather than running around at random. Therefore there is no friction. The liquid is a superfluid.

That makes for simplification, but it also presents a problem. To investigate turbulence, the natural place to start is by imagining a rotating bucket of superfluid. In fact, a superfluid cannot be set into simple rotation. To do so, we would have to force it into 'solid body rotation', for which the tangential velocity increases linearly from the centre. But there is no way in which we can devise a wavefunction for the superfluid, describing the state of every constituent particle, that will achieve this. It is easy to create a wavefunction in which the velocity falls as we go